IDAP INSTITUTE OF DEFENSE ACQUISITION PROGRAM

JAMS JOURNAL OF ADVANCES IN MILITARY STUDIES

2026, Vol. 9, No. 2, 119-141. https://doi.org/10.37944/advmil.9.2.202610.006

A Study on Improving the Defense Software Development Cost Estimation System

Lee Jisun* · Hwang Hyunho** · Choi Eunbi***

ABSTRACT

Defense systems increasingly rely on software (SW) and artificial intelligence (AI), yet defense SW development costs continue to be estimated primarily based on labor effort and wage rates, with insufficient consideration of project characteristics and technical complexity. This study examined practitioners' perceptions of the current cost estimation system and explored potential improvements. Survey items were developed through a preliminary survey of 35 small and medium-sized defense SW firms, followed by a five-point Likert-scale survey of 96 practitioners. All eight identified limitations received mean scores significantly above the neutral midpoint ($M = 3.71$-4.00, $p < .001$). Across all 13 cost items, perceived necessity significantly exceeded current levels of cost recognition ($d = 0.36$-1.33), with the largest gaps (+1.64 to +1.78) observed in maintenance and performance enhancement, AI testing and validation, dataset construction, and model training. AI-related items were perceived as equally necessary as non-AI items but were significantly less reflected in current cost estimates. Respondents also expressed strong support for project-specific adjustments ($M = 4.12$) and five proposed adjustment factors ($M = 3.84$-4.08). These findings suggest that effort-based estimation should be retained as the baseline while AI development and life-cycle costs are separately identified and project-specific adjustment factors are incorporated. The proposed cost estimation framework remains conceptual and requires further validation using actual project cost data.

Keywords : Defense Software Development Cost, Software Cost Estimation, AI Development Cost, Life-Cycle Cost, Project-Specific Adjustment Factors

* (First Author, Corresponding Author) Korea Defense Technology Society (KIDET), jisun0319@gmail.com, https://orcid.org/0000-0003-2345-7698

** (Co-author) Department of Defense Convergence Technology, Hanyang Cyber University, 1250002@hycu.ac.kr, https://orcid.org/0009-0001-4065-0943

*** (Co-author) Korea Research Institute for Defense Technology Planning and Advancement (KRIT), choieunbi@krit.re.kr

www.kci.go.kr

Ⅰ. 서론

최근 국방 분야는 하드웨어 중심의 무기체계에서 소프트웨어(SW)와 인공지능(AI)을 기반으로 한 첨단체계로 빠르게 전환되고 있다. 첨단 무기체계에서 SW는 핵심 기능의 구현뿐만 아니라 운용 과정에서의 지속적인 성능개선을 좌우하는 핵심 요소이다. 미국 국방부는 이러한 특성을 반영하여 신속하고 반복적인 SW 개발과 제공을 별도의 획득경로로 제도화하였다(U.S. Department of Defense, 2020).[1] 특히 AI 기반 체계는 데이터 구축, 모델 학습, 시험・검증 및 재학습이 반복적으로 이루어지며, 전력화 이후에도 운용 환경과 요구사항의 변화에 따라 지속적인 유지보수와 성능개선이 필요하다.

이러한 개발 방식의 변화에도 불구하고 국방 SW 개발비는 공공 SW 대가산정체계와 방산원가체계가 교차 적용되는 제도적 환경에서 산정되고 있다(Korea Artificial Intelligence and Software Industry Association, 2025; Ministry of National Defense, 2026). 현행 투입공수법은 투입 인력의 등급별 노임단가와 투입 기간을 기준으로 직접인건비를 산정하고 직접경비를 더하는 방식이다. 이 방식은 비용 산정이 비교적 간편하다는 장점이 있으나, 개발 결과물의 기능적 가치, 전문 인력의 생산성 차이, 기업이 축적한 기술 자산과 재사용 자산의 가치, 개발 과정의 불확실성을 충분히 반영하는 데 한계가 있다. 또한 무기체계 사업에 적용되는 방산원가 기준과 공공 SW 사업에 적용되는 「SW사업 대가산정 가이드」의 차이로 인해 사업이나 기관에 따라 인정되는 비용의 범위와 수준이 달라질 가능성도 제기되고 있다.[2]

국방 SW 사업에는 폐쇄망 개발 환경과 보안 규정 준수, 국방 표준 및 기존 체계와의 연동, 시험평가・품질보증, 운용 환경 실증 등 일반 SW 사업과 구별되는 비용 요소가 존재한다. AI 기반 사업의 경우에는 데이터 수집・정제・구축, 모델 학습, 알고리즘 조정, 반복적인 시험・검증을 위한 추가 비용도 발생한다. 아울러 전력화 이후의 유지보수와 성능개선은 무기체계의 운용성과 신뢰성을 유지하기 위한 핵심 활동이지만, 이러한 비용 요소가 현행 산정체계에 어느 정도 반영되고 있는지, 현장에서는 그 필요성을 어떻게 인식하고 있는지에 대한 실증적 연구는 제한적이다.

이에 본 연구는 국방 SW 사업 실무자를 대상으로 설문조사를 실시하여 현행 개발비 산정체계에 대한 현장 인식을 실증적으로 진단하고, 국방 SW의 특성을 반영한 개발비 산정체계의 개선방향을 도출한다. 이를 위해 첫째, 현행 투입공수 기반 산정체계의 제한사항에 대한 실무자의 인식을 분석한다. 둘째, 국방 SW 사업의 주요 비용 항목별로 현행 반영 수준과 반영 필요성을 비교하여 그 격차를 확인하고 우선적으로 개선해야 할 항목을 도출한다. 셋째, 사업 특성에 따른 추가 보정의 필요성과 제시된 보정계수의 적정성을 평가한다. 이상의 분석 결과를 종합하여 국방 SW 개발비 산정체계가 갖추어야 할 기본원칙과 구체적인 제도 개선방향을 제시한다.

---

1) U.S. Department of Defense. (2020, October 2). Operation of the Software Acquisition Pathway (DoD Instruction 5000.87).

2) 한국인공지능・소프트웨어산업협회. (2025). SW사업 대가산정 가이드(2025년 개정판).
국방부. (2026). 방산원가대상물자의 원가계산에 관한 규칙(국방부령 제1202호). 국가법령정보센터.

www.kci.go.kr

본 논문의 구성은 다음과 같다. II장에서는 SW 개발비 산정 방식과 국방 SW의 특수성에 관한 이론적 배경 및 선행연구를 검토한다. III장에서는 설문 문항의 도출 과정과 조사 설계 및 분석 방법을 제시하고, IV장에서는 현행 산정체계의 제한사항, 비용 항목별 반영 수준과 필요성의 격차 및 보정계수에 대한 분석 결과를 제시한다. 마지막으로 V장에서는 연구 결과를 종합하여 국방 SW 개발비 산정체계의 개선방향을 논의하고, 연구의 한계와 향후 연구과제를 제시한다.

Ⅱ. 관련 연구

SW 개발비 산정 방식은 크게 투입공수 기반 방식과 기능규모 기반 방식으로 구분된다. 투입공수법은 개발에 필요한 인력의 등급과 인원, 투입기간을 산정하고 여기에 기준 노임단가를 적용하여 직접인건비를 구한 뒤 제경비, 기술료, 직접경비를 가산하는 방식으로, 산정 근거가 명확하고 관리가 용이하다는 장점이 있다. 반면 기능점수(Function Point) 방식은 SW가 제공하는 기능의 규모를 측정하여 기능점수당 단가를 적용하는 방식으로, 투입 인력이 아닌 산출물 규모를 기준으로 하기 때문에 생산성 차이가 대가에 반영될 수 있다. 국내 공공 SW 사업에서는 「SW사업 대가산정 가이드」에 따라 기능점수 방식을 원칙으로 하되 기능점수 산정이 어려운 경우 투입공수 방식을 보완적으로 적용하도록 하고 있다(Korea Artificial Intelligence and Software Industry Association, 2025).[3)]

SW 비용 추정 분야에서는 규모와 비용동인을 입력으로 공수와 일정을 추정하는 매개변수 모형이 오래전부터 개발되어 왔다. COCOMO II는 원시코드 라인 수나 기능점수로 측정한 규모에 규모인자(scale factor)와 노력승수(effort multiplier)를 적용하여 공수와 일정을 추정하며(Boehm et al., 2000),[4)] PRICE-S와 SEER-SEM 역시 규모, 개발 환경, 운용 플랫폼, 신뢰도 요구 등 다수의 매개변수로 비용과 일정을 산출하는 모형으로 국방 획득사업의 비용분석에 활용되어 왔다(Park, 1988; Galorath & Evans, 2006).[5)] 기능점수 방식은 Albrecht(1979)가 산출물 기준의 규모 측정 방법으로 제안한 이후 표준화되었으며,[6)] 국내 「SW사업 대가산정 가이드」의 기능점수 방식도 이에 기반한다. 이러한 추정모형 연구에서는 모형 간 정확도 비교와 실적 자료를 통한 검증이 주요 연구과제로 다루어져 왔다

3) 한국인공지능·소프트웨어산업협회. (2025). SW사업 대가산정 가이드(2025년 개정판).

4) Boehm, B. W., Abts, C., Brown, A. W., Chulani, S., Clark, B. K., Horowitz, E., Madachy, R., Reifer, D. J., & Steece, B. (2000). Software cost estimation with COCOMO II. Prentice Hall.

5) Park, R. E. (1988). The central equations of the PRICE software cost model. Paper presented at the 4th COCOMO Users' Group Meeting, Software Engineering Institute, Carnegie Mellon University.
Galorath, D. D., & Evans, M. W. (2006). Software sizing, estimation, and risk management: When performance is measured performance improves. Auerbach Publications.

6) Albrecht, A. J. (1979). Measuring application development productivity. In Proceedings of the Joint SHARE/GUIDE/IBM Application Development Symposium (pp. 83-92).

www.kci.go.kr

(Jørgensen & Shepperd, 2007).[7]

국방 분야에서도 이러한 모형을 국방 환경에 맞게 적용・보완하려는 연구가 이루어졌다. 미국에서는 국방부 SW 자원자료(SRDR)를 바탕으로 국방 SW의 응용영역별 생산성 기준과 비용 추정 관계식이 제시되었고(Clark & Madachy, 2015),[8] 최근에는 국방부 애자일 사업 자료를 바탕으로 초기 요구사항 수를 규모 척도로 하는 공수・일정 추정모형이 제안되었다(Rosa et al., 2022).[9] 국내에서는 국방 SW 비용산정에 적용되던 PRICE-S를 한국소프트웨어산업협회 모델과 비교하여 보완하고 기능점수 백파이어링 절차를 제시한 연구(Lee, B. E., & Kang, S. J., 2005)[10]와, COCOMO II, PRICE-S, SEER-SEM 등 범용 모형이 국방 환경의 특성을 반영하지 못한다는 점에 주목하여 무기체계 내장형 SW에 특화된 비용 추정 모형의 개발 절차를 제시한 연구(Kwon, T. Y., 2010)[11]가 있다. 한편 AI 기반 SW는 데이터 의존성, 재학습, 운용 모니터링 등으로 인해 배포 이후에도 유지비용이 누적되는 특성이 있다는 점이 지적되었다(Sculley et al., 2015).[12]

국방 분야에서는 무기체계 연구개발 사업에 방산원가 기준이 적용되는 경우가 많아, SW 개발활동이라 하더라도 제조업 중심의 원가계산 체계에 따라 제비율과 이윤이 산정된다. 따라서 동일한 SW 개발활동이라도 사업의 법적 성격과 적용 기준에 따라 인정되는 비용 항목과 이윤 산정방식이 달라질 수 있다(Ministry of National Defense, 2026).[13]

국방 SW는 신뢰성과 안전성 요구가 높아 개발 과정에서 정적・동적 분석을 포함한 SW 신뢰성 시험, 상호운용성 평가, 보안성 검증 등 다단계 시험・검증 절차를 거치며, 개발관리 규정에 따라 단계별 산출물(SDRL)을 작성・검토하고 형상관리를 수행하여야 한다(Defense Acquisition Program Administration, 2024).[14] 또한 보안 규정에 따른 폐쇄망・망분리 환경은 오픈소스, 클라우드, 외부 개발도구 활용을 제한하여 개발 생산성에 영향을 미친다. 따라서 이러한 의무활동이 작업분류체계와

7) Jørgensen, M., & Shepperd, M. (2007). A systematic review of software development cost estimation studies. IEEE Transactions on Software Engineering, 33(1), 33-53. https://doi.org/10.1109/TSE.2007.256943

8) Clark, B., & Madachy, R. (Eds.). (2015). Software cost estimation metrics manual for defense systems. Software Metrics Inc.

9) Rosa, W., Clark, B. K., Madachy, R., & Boehm, B. W. (2022). Empirical effort and schedule estimation models for agile processes in the US DoD. IEEE Transactions on Software Engineering, 48(8), 3117-3130. https://doi.org/10.1109/TSE.2021.3080666

10) 이병은, 강성진. (2005). 백파이어링을 이용한 군사용 소프트웨어 초기단계 개발비용 산정 기법. 정보처리학회논문지D, 12D(5), 737-744.

11) 권태용. (2010). 국방 환경을 고려한 무기체계 내장형 소프트웨어 비용 추정 모델 개발 [석사학위논문]. 한국과학기술원.

12) Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. Advances in Neural Information Processing Systems, 28, 2503-2511.

13) 국방부. (2026). 방산원가대상물자의 원가계산에 관한 규칙(국방부령 제1202호). 국가법령정보센터.

14) 방위사업청. (2024). 무기체계 소프트웨어 개발 및 관리 매뉴얼(방위사업청 매뉴얼 제2024-6호).

www.kci.go.kr

원가항목에 명시되지 않으면 실제 수행공수와 산정공수 사이에 구조적 차이가 발생할 수 있다.

AI SW 개발은 데이터의 규모와 품질이 성능을 좌우하며, 요구되는 성능을 달성할 때까지 데이터 보완, 모델 학습, 검증, 재학습을 반복하는 탐색적 특성을 갖는다. 이러한 활동은 전통적 코드 구현 공수와 성격이 다르므로 데이터 구축・정제・라벨링, 합성데이터 생성, 학습 인프라 및 도구 비용을 산정표에서 구분할 필요가 있다. 이는 AI 위험관리를 설계부터 운용과 모니터링까지 전 생애주기에 걸쳐 수행해야 한다는 NIST(2023)[15]의 관점과도 부합한다. 애자일 개발 역시 반복적・점진적 개발과 요구사항의 지속적 조정을 전제로 하므로, 확정형 요구사항(ROC)과 고정형 계약을 기반으로 하는 전통적 획득체계와 조화되기 어렵다는 논의가 국내외에서 이루어지고 있다(U.S. Department of Defense, 2020).[16]

이상의 추정모형 연구는 주로 과거 사업의 실적 자료를 바탕으로 공수와 비용을 추정하는 모형을 개발하거나 그 정확도를 높이는 데 초점을 두었다. 그러나 국내 국방 SW 사업에서는 추정된 공수가 방산원가 기준이나 「SW사업 대가산정 가이드」에 따라 어떤 비용 항목으로 얼마나 인정되는지가 실제 대가를 좌우하며, 이러한 비용 인정의 문제는 추정모형 연구에서 상대적으로 다루어지지 않았다. 또한, 선행연구는 투입공수법의 구조적 한계, 국방 SW 특수비용의 존재, AI・애자일 개발과 기존 산정체계 간의 부정합을 각각 지적하고 있으나, 이러한 쟁점들에 대해 현장 실무자가 실제로 어느 정도 공감하는지, 비용 항목별로 현재 어느 정도 반영되고 있으며 무엇이 우선적으로 보완되어야 하는지를 실증적으로 규명한 연구는 부족하다. 본 연구는 비용 추정모형을 새로 개발하는 것이 아니라 현행 비용 인정체계에 대한 현장 인식을 진단한다는 점, 그리고 개발 대상 SW의 규모와 비용동인 대신 비용 항목별 반영 수준과 필요성의 격차를 측정하여 제도 보완의 우선순위를 도출한다는 점에서 기존 연구와 구별된다. 본 연구는 정성적 조사와 정량적 조사를 순차적으로 결합하여 이 공백을 메우고자 하며, 앞서 서론에서 제시한 세 가지 연구문제를 중심으로 분석을 수행한다. 본 연구는 탐색적 정책연구의 성격을 고려하여 별도의 방향가설을 확정하기보다, 각 연구문제에 대응하는 통계검정을 통해 인식의 방향과 차이의 크기를 확인하였다.

Ⅲ. 연구방법

3.1 연구 설계

본 연구는 국방 SW 사업 실무자를 대상으로 한 구조화 설문조사를 중심으로 하되, 설문 문항의 내

15) Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

16) U.S. Department of Defense. (2020, October 2). Operation of the Software Acquisition Pathway (DoD Instruction 5000.87).

www.kci.go.kr

용타당성을 확보하기 위해 개방형 예비조사를 선행하는 탐색적 순차 혼합설계(exploratory sequential mixed design)를 채택하였다(Creswell & Plano Clark, 2018).[17] 예비조사(1차)에서는 국방 SW 사업 참여 중소기업을 대상으로 현행 산정체계의 문제점과 비용 항목을 개방형으로 발굴하고, 본조사(2차)에서는 이를 구조화된 문항으로 전환하여 정량적으로 검증하였다. 연구 절차의 개요는 <Table 1>과 같다.

〈Table 1〉 Overview of the Research Procedure

Phase	Respondents & Period	Survey Content	Analysis Method
Preliminary Survey (1st)	35 SMEs participating in defense SW projects (Apr. 20–30, 2026)	Project areas, challenges in man-hour-based cost estimation, defense-specific additional costs, barriers to technology adoption, and additional comments (five open-ended questions)	Inductive categorization of responses → derivation of main survey items
Main Survey (2nd)	96 practitioners from companies participating in defense SW projects (Aug. 12–30, 2026; online and offline at a public hearing)	Agreement with eight system limitations, current implementation levels and perceived necessity of 13 cost items, the need for additional adjustments, the adequacy of five adjustment factors, and open-ended comments (five-point Likert scale)	Descriptive statistics, one-sample and paired t-tests, effect size, gap analysis, IPA, AI vs. non-AI cluster comparison, comparison by SW type (ANOVA)

3.2 예비조사 및 본조사 개요

예비조사는 2026년 4월 20일부터 30일까지 국방 SW 사업에 참여하는 중소기업 35개사를 대상으로 실시하였다. 응답 기업의 주요 참여 사업 영역은 무기체계 SW 19개사(54.3%), M&S 9개사(25.7%), 전장관리정보체계 5개사(14.3%), 사이버 체계 5개사(14.3%), 기타 8개사(22.9%)였다(복수응답). 응답 내용을 귀납적으로 범주화한 결과, 투입공수법의 한계는 고급인력・자체투자비 미보전, 획일적 등급체계, 생산성이 높을수록 불리한 구조, 사전 공수 예측의 한계, 원가기준 적용 편차, 부수업무 공수 미반영의 여섯 가지로, 국방 특수비용은 시험・검증, 문서화・산출물, 보안환경, 라이선스・데이터, 형상・변경관리의 다섯 가지로 정리되었으며, 이를 토대로 본조사의 제한사항 8개 항목과 비용 항목 13개를 구성하였다. 예비조사 결과는 모집단의 비율을 추정하는 자료가 아니라 본조사 문항의 내용영역을 식별하는 질적 자료로 활용하였다.

본 조사는 2026년 8월 12일부터 30일까지 국방 SW 사업 참여기업 실무자를 대상으로 설문을 진행하여 총 96명의 응답을 수집하였다. 응답자가 문제를 가장 많이 경험한 SW 유형은 기본형(일반 커스터마이징) 40명(41.7%), 모델형(모델 구현・학습 중심) 22명(22.9%), 사업별 상이 20명(20.8%),

17) Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

www.kci.go.kr

데이터형(데이터 구축 중심) 14명(14.6%)이었다. 본조사에서는 기업 규모・직무・경력 등 개인 배경 변수를 수집하지 않았으므로 표본 특성의 기술과 집단 비교는 주 참여 SW 유형으로 한정하였다. 이로 인해 응답자가 중소 전문기업과 체계업체 중 어느 쪽에 분포하는지, 직무와 경력에 따라 응답이 달라지는지를 확인할 수 없어 표본의 대표성에는 제약이 있다. 또한 응답자는 모두 SW 개발을 수행하는 공급자 측 실무자이므로, 비용 인정수준에 대한 응답에는 비용이 더 인정되기를 기대하는 이해관계가 반영될 수 있다. 따라서 본 연구의 결과는 실제 비용 인정수준이 아니라 공급자 측 실무자의 인식으로 해석하였으며, 이러한 편향 가능성과 후속 검증 방향을 제시하였다. 측정도구는 네 영역으로 구성하였다. 첫째, 현행 산정방식의 제한사항 8개 항목(SW 개발 특성과 맞지 않는 방식, 고급 SW 인력 가치 반영 부족, 자체 투자・기술자산 보상 부족, 국방 분야 특수비용 미반영, 실 투입비용 대비 보전 부족, SW 결과물 가치 반영 부족, 양산・후속 단계 SW 가치 미인정, 행정・제도상의 어려움)에 대한 공감도이다. 둘째, 국방 SW 사업에서 발생하는 13개 비용 항목에 대한 현재 개발비 반영수준과 사업 수행 시 필요성이다. 셋째, 사업 특성에 따른 기술료율 차등(20～40%) 등 추가 보정의 필요성과 5개 보정계수(규모, 연계 복잡성, 성능 요구수준, 운영환경 호환성, 보안성 요구수준)의 적정성이다. 넷째, 추가 고려 비용 항목, 추가 보정요소, 개발비 산정 전반에 대한 개방형 의견이다. 정량 문항은 5점 리커트 척도(1=매우 그렇지 않다, 5=매우 그렇다)로 측정하였으며, 무응답은 결측으로 처리하여 해당 문항의 통계 산출에서 제외하였다(문항별 유효응답 94～96명).

3.3 분석방법

수집된 자료는 문항별 평균, 표준편차, 응답 분포를 산출하는 기초통계 분석을 실시하였으며, 4～5점 응답 비율을 긍정응답률(공감률・필요율・적정률), 1～2점 응답 비율을 부정응답률(미반영률)로 정의하였다. 제한사항 및 보정계수 문항은 척도 중립값(3점)을 기준으로 일표본 t검정을 실시하여 공감・수용 수준의 통계적 유의성을 확인하였다. 13개 비용 항목에 대해서는 필요성 평균에서 현재 반영수준 평균을 차감한 격차를 산출하고 격차가 큰 순서로 제도 보완의 우선순위를 부여하였으며, 격차의 유의성은 대응표본 t검정으로 검정하고 효과크기(Cohen’s d)를 함께 제시하였다. 항목 평균에 기반한 분석이 개별 응답자의 인식과 괴리되지 않는지 확인하기 위해 응답자별 평균 격차의 분포도 검토하였다. 격차 분석의 한계를 보완하기 위해 필요성과 반영수준의 중앙값(각각 4.04, 2.63)을 기준축으로 하는 중요도-반영도 분석(IPA)을 실시하여 각 항목을 4개 사분면에 배치하였다(Martilla & James, 1977).[18] 또한 AI 관련 4개 항목과 비AI 9개 항목을 클러스터로 구분하여 응답자별 클러스터 평균을 대응표본 t검정으로 비교하고, SW 유형별로 응답자별 평균 격차의 차이를 일원분산분석(ANOVA)으로 검정하였다. 제한사항, 비용 항목, 보정계수 문항은 서로 다른 비용요소를 개별적으로

18) Martilla, J. A., & James, J. C. (1977). Importance-performance analysis. Journal of Marketing, 41(1), 77-79. https://doi.org/10.2307/1250495

www.kci.go.kr

평가하는 형성적(formative) 지표로 설계되었으므로 항목 간 높은 상관을 전제하지 않고 각 비용요소를 개별적으로 해석하였다(Bollen & Lennox, 1991; Diamantopoulos & Winklhofer, 2001).[19]

Ⅳ. 분석결과

4.1 현행 산정체계의 제한사항에 대한 공감도

현행 산정방식의 제한사항 8개 항목에 대한 공감도를 조사한 결과, 모든 항목의 평균이 3.71 이상, 긍정응답률이 59.4% 이상으로 나타났으며, 일표본 t검정 결과 8개 항목 모두 중립값 3점보다 통계적으로 유의하게 높았다(t=7.39～13.51, p<.001)(Table 2). '자체 투자・기술자산 보상 부족'이 평균 4.00, 긍정응답률 82.3%로 가장 높았고 표준편차(0.73)도 가장 낮아 응답자 간 의견 일치도가 높았다. '양산・후속 단계 SW 가치 미인정'(3.97)과 '국방 분야 특수비용 미반영'(3.94)이 뒤를 이었으나, 전자는 표준편차가 1.09로 가장 커 체감 정도에 응답자 간 차이가 있는 것으로 나타났다. 각 항목에 대한 부정응답은 4.2～7.3%에 그쳤으며, 'SW 개발 특성과 맞지 않는 방식'은 '보통' 응답 비율이 36.5%로 가장 높아 현행 방식과 SW 개발 특성 간 불일치에 대한 체감 수준이 응답자별로 상이함을 보여준다.

〈Table 2〉 Agreement with the Limitations of the Current Cost Estimation System
(N = 96, One-Sample t-Test, Test Value = 3)

Item	Mean	SD	Positive Response Rate (4-5 pts)	t	p
Insufficient compensation for in-house investment and technology assets	4.00	0.73	82.3%	13.51	<.001
SW value not recognized in production and follow-on phases	3.97	1.09	69.8%	8.71	<.001
Defense-specific costs not reflected	3.94	0.92	74.0%	10.03	<.001
Insufficient reflection of the value of senior SW personnel	3.81	1.01	66.7%	7.89	<.001
Insufficient reflection of the value of SW deliverables	3.81	0.79	66.7%	10.13	<.001
Method ill-suited to the nature of SW development	3.78	0.85	59.4%	9.02	<.001

19) Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. Psychological Bulletin, 110(2), 305-314. https://doi.org/10.1037/0033-2909.110.2.305
Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. Journal of Marketing Research, 38(2), 269-277. https://doi.org/10.1509/jmkr.38.2.269.18845

www.kci.go.kr

Item	Mean	SD	Positive Response Rate (4-5 pts)	t	p
Insufficient reimbursement relative to actual costs incurred	3.75	0.93	69.8%	7.91	<.001
Administrative and institutional difficulties	3.71	0.94	62.5%	7.39	<.001

4.2 비용 항목별 반영수준, 필요성 및 격차

13개 비용 항목의 현재 반영수준은 모든 항목에서 평균 3.33 이하로 나타났으며, 특히 AI 관련 4개 항목과 유지보수・성능개선은 평균 2.36～2.48, 미반영률 55.2～61.7%로 절반 이상의 응답자가 충분히 반영되지 않고 있다고 인식하였다. 반영수준에서 1점(전혀 반영되지 않음) 응답은 요구사항 변경 대응(22명), 유지보수・성능개선(19명), 클라우드・라이선스(19명)에서 많았고, 국방 표준・프로세스 준수는 1점 응답이 없는 유일한 항목이었다. 반면 필요성은 모든 항목에서 평균 3.78 이상, 필요율 66.7% 이상으로 높게 나타났으며, '시험・검증・품질보증'과 '유지보수・성능개선'이 각각 4.26으로 가장 높았다. 필요성 평균에서 반영수준 평균을 차감한 격차는 13개 항목 모두 양(+)의 값을 보였고 +0.45에서 +1.78까지 분포하였으며, 대응표본 t검정 결과 모든 항목에서 통계적으로 유의하였다($p<.001$). 효과크기는 국방 표준・프로세스 준수($d=0.36$)를 제외한 12개 항목에서 큰 수준($d\geq 0.74$)이었고, 격차 상위 4개 항목은 $d=1.16$～1.33으로 매우 컸다(Table 3).

〈Table 3〉 Levels of Cost Item Inclusion, Perceived Necessity, and Gaps
(N = 96; Inclusion Level: n = 94; Gap = Necessity−inclusion Level; All t-Tests: p 〈 .001)

Cost Item	Refl. Mean (SD)	Non-refl. Rate	Need Mean (SD)	Need Rate	Gap	t	d	Rank
Maintenance & performance improvement	2.48 (1.09)	55.2%	4.26 (0.91)	85.4%	+1.78	12.02	1.23	1
AI testing & verification	2.41 (0.97)	59.4%	4.19 (0.87)	78.1%	+1.78	12.88	1.31	2
AI data construction, synthesis & preprocessing[a]	2.36 (0.95)	61.7%	4.04 (0.91)	78.1%	+1.68	12.92	1.33	3
AI model training	2.41 (0.97)	59.4%	4.04 (0.94)	74.0%	+1.64	11.32	1.16	4
Responding to requirement changes	2.62 (1.15)	40.6%	4.19 (0.87)	78.1%	+1.56	11.13	1.14	5
Cloud fees & licenses	2.62 (1.17)	47.9%	4.19 (0.84)	81.2%	+1.56	9.76	1.00	6
AI algorithm tuning	2.48 (1.01)	55.2%	3.96 (0.86)	78.1%	+1.48	11.61	1.19	7
Security environment setup	2.56 (1.05)	47.9%	3.96 (0.89)	74.0%	+1.40	9.70	0.99	8
Testing, verification & quality assurance	3.00 (1.13)	37.5%	4.26 (0.77)	88.5%	+1.26	8.72	0.89	9
Development environment &	2.81 (1.11)	42.7%	4.00 (0.78)	78.1%	+1.19	9.06	0.92	10

www.kci.go.kr

Cost Item	Refl. Mean (SD)	Non-refl. Rate	Need Mean (SD)	Need Rate	Gap	t	d	Rank
infrastructure setup								
Documentation & deliverable preparation	2.84 (1.13)	40.6%	4.00 (0.92)	81.2%	+1.16	7.24	0.74	11
Logistics support analysis (LSA) & IPS development	2.84 (1.03)	37.5%	3.90 (0.75)	66.7%	+1.05	8.35	0.85	12
Compliance with defense standards & processes	3.33 (0.83)	18.8%	3.78 (0.93)	74.0%	+0.45	3.52	0.36	13

4.3 AI · 비AI 항목 클러스터 비교

AI 관련 항목이 격차 상위권에 집중된 원인을 확인하기 위해 AI 관련 4개 항목과 비AI 9개 항목의 필요성과 반영수준을 비교하였다(Table 4). 두 클러스터의 필요성 평균은 4.06으로 동일한 반면, 반영수준은 AI 항목이 2.41로 비AI 항목(2.79)보다 0.38점 낮았다. 응답자별 클러스터 평균을 대응표본 t검정으로 비교한 결과, 반영수준의 차이는 통계적으로 유의하였으나($t(95)=-5.31$, $p<.001$) 필요성의 차이는 유의하지 않았다($t(95)=-0.03$, $p=.973$). 따라서 AI 관련 항목의 큰 격차(+1.64 대 +1.26)는 필요성이 특별히 높게 평가되었기 때문이 아니라 현행 산정체계에서의 반영수준이 낮기 때문임을 확인할 수 있다. SW 유형별 AI 항목의 평균 격차는 '사업별 상이' 유형이 2.03으로 가장 컸고 기본형과 데이터형이 각각 1.62, 모델형이 1.36이었다.

〈Table 4〉 Comparison of AI and Non-AI Cost Item Clusters

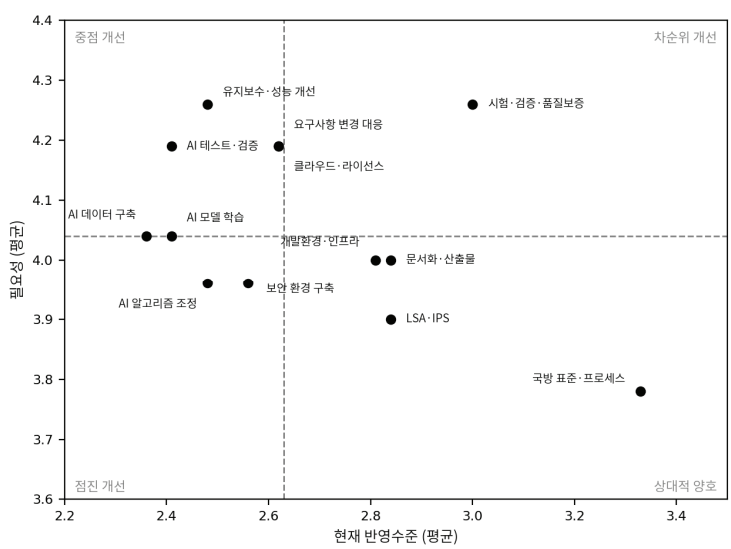
Category	Necessity	Reflection Level	Gap	Difference Between Clusters
4 AI-related items	4.06	2.41	+1.64	Reflection level: $t=-5.31$, $p<.001$
9 non-AI items	4.06	2.79	+1.26	Necessity: $t=-0.03$, $p=.973$

4.4 중요도-반영도 분석(IPA)

필요성 중앙값 4.04와 반영수준 중앙값 2.63을 기준축으로 13개 항목을 사분면에 배치한 결과는 <Figure 1> 및 <Table 5>와 같다. 사분면은 기준값 이상을 '高', 미만을 '低'로 구분하는 규칙에 따라 분류하였으며, 필요성이 기준값과 같은(4.04) 'AI 데이터 구축'과 'AI 모델 학습'은 필요성 高로 분류하였다. 이 기준에 따르면 필요성이 높고 반영 수준이 낮은 중점 개선 영역에는 '유지보수 · 성능개선', 'AI 데이터 구축', 'AI 모델 학습', 'AI 테스트 · 검증', '요구사항 변경 대응', '클라우드 · 라이선스'의 6개 항목이 위치하며(<Figure 1>), 이 중 앞의 4개 항목은 격차 분석의 상위 1~4위 항목과 일치한다. 반면 '요구사항 변경 대응'과 '클라우드 · 라이선스'는 반영 수준(2.62)이 기준축(2.63)과 0.01

www.kci.go.kr

점 이내로 근접하여, 기준축을 소폭만 조정하여도(예: 반올림한 항목 평균의 중앙값 2.62 적용) 차순위 개선 영역으로 이동할 수 있는 경계 항목이며, 격차 분석에서도 5~6위로 상위 4개 항목과 구분된다. 이에 <Table 5>에서는 사분면에 따른 분류(중점 개선)는 그대로 유지하되, 두 항목의 분류가 기준축 설정에 민감하다는 저자의 판단을 반영하여 '중점 개선(경계 항목)'으로 구분하여 표기하였다. 따라서 본 연구에서 '격차 분석과 IPA에서 공통으로 도출된 상위 4개 항목'은 IPA 중점 개선 영역 6개 항목 중 경계 항목 2개를 제외한 항목을 의미하며, 경계 항목 2개도 사분면상 중점 개선 영역에 속하므로 실질적 우선순위는 상위 4개 항목에 준하여 해석한다. '시험 · 검증 · 품질보증'은 필요성(4.26)과 반영 수준(3.00)이 모두 기준값 이상인 차순위 개선 영역에 위치하여, 일정 수준 반영되고 있으나 필요성과의 격차(+1.26)가 커 적용 기준의 일관성 확보가 필요한 항목으로 해석된다. '유지보수 · 성능개선', 'AI 데이터 구축', 'AI 모델 학습', 'AI 테스트 · 검증'은 필요성이 높고 반영수준은 낮은 중점 개선 영역에 위치하였으며, 이는 격차 분석의 상위 4개 항목과 일치하여 두 분석방법 간 결과의 일관성을 보여준다. '요구사항 변경 대응', '클라우드 · 라이선스', '시험 · 검증 · 품질보증'은 반영수준이 상대적으로 높더라도 필요성과의 격차가 커 차순위 개선 영역으로, '보안 환경 구축'과 'AI 알고리즘 조정'은 필요성 · 반영수준 모두 중앙값 이하이나 격차가 작지 않아 점진 개선 영역으로 분류되었다. '요구사항 변경 대응'과 '클라우드 · 라이선스'는 반영수준(2.62)이 중앙값(2.63)에 근접한 경계 항목으로, 기준축 설정에 따라 중점 개선 영역으로 분류될 수 있어 실질적 우선순위는 중점 개선 항목에 준하여 해석할 필요가 있다. 또한 IPA 사분면은 표본 중앙값을 기준으로 한 상대적 분류이므로, 축의 위치를 달리한 민감도 분석과 실제 원가자료를 함께 검토한 뒤 제도 우선순위로 확정할 필요가 있다.



〈Figure 1〉 Results of the Importance-Performance Analysis of Cost Items (Dashed Lines: Median Values of Necessity and Inclusion Level)

www.kci.go.kr

〈Table 5〉 Classification of Cost Items by IPA Quadrant

Quadrant	Item	Interpretation
Priority Improvement (High necessity, low reflection)	Maintenance & performance improvement, AI data construction, AI model training, AI testing & verification	Institutional improvement needed as a priority(consistent with gap ranks 1-4)
Priority Improvement - Borderline Items(High necessity, low reflection)	Responding to requirement changes, cloud & licenses	Positioned within the Priority Improvement quadrant but marked separately due to its proximity to the reference axis; the same improvement priority is maintained
Secondary Improvement (High necessity, high reflection)	Testing, verification & quality assurance	Reflected in the current system, but a substantial gap remains; greater consistency in application criteria is required
Gradual Improvement (Medium or lower necessity and reflection)	Security environment setup, AI algorithm tuning	Detailed application criteria by project characteristics needed
Relatively Satisfactory (Low necessity, high reflection)	Documentation & deliverables, development environment & infrastructure, LSA & IPS, defense standards & processes	Relatively low priority, but continuous monitoring needed

4.5 개발비 보정체계에 대한 수용성

사업 특성에 따른 기술료율 차등 등 추가 보정의 필요성은 평균 4.12(SD 1.07), 긍정응답률 74.0%로 높게 나타났다(일표본 t검정 $p<.001$). 다만 부정응답이 9.4%로 다른 문항보다 다소 높아, 보정의 방향에는 공감하되 획일적 요율 차등에는 유보적인 응답자가 일부 존재함을 시사한다. 5개 보정계수의 적정성은 모든 항목에서 평균 3.84 이상으로 긍정적으로 평가되었으며(모두 $p<.001$), '성능 요구수준'(4.08)이 가장 높았고 '운영환경 호환성'(4.04), '연계 복잡성'(4.03), '규모'(3.95), '보안성 요구수준'(3.84) 순이었다(Table 6). 부정응답은 항목별 3.1～5.2%에 그쳤다. '보안성 요구수준'은 평균이 가장 낮고 표준편차(0.91)가 상대적으로 커, 시큐어 코딩・취약점 점검・암호화 등 다양한 보안 요구를 하나의 보정계수로 적용하는 방식과 보정수준에 대해 응답자 간 판단이 엇갈린 것으로 해석된다. 이는 개방형 응답에서 '보안성 요구수준을 하나의 계수로 묶는 것은 무리가 있어 세부 기준이 필요하다'는 의견이 제시된 것과 일치한다.

보정계수 외에 추가로 고려할 요소를 묻는 개방형 문항에서는 무기체계 등급 및 임무 중요도, 실시간・안전필수(safety-critical) 요구 여부, SW 신뢰성 시험 수준 등 시험검증 강도, AI 모델의 재학습・지속 고도화 필요 여부, 폐쇄망 개발 여부(개발환경 제약 정도), 연계 대상 체계의 문서화 수준 및 인터페이스 안정성, RFP 이후 요구사항 변경 빈도, 체계 규모 대비 개발기간 압축 정도, 국산화・독자기술 개발 여부가 제시되었다. 이들 요소는 기술적 복잡성 중심으로 구성된 현행 5개

www.kci.go.kr

보정계수에 임무 중요도, 검증 강도, 개발환경 제약, 일정 · 요구 변동성 등 국방 사업의 운용적 특성을 보완하는 방향으로 검토할 수 있다.

〈Table 6〉 Evaluation of the Appropriateness of Adjustment Factors
(One-Sample t-Test, Test Value = 3; All p 〈 .001)

Adjustment Factor	Mean	SD	Adequacy Response Rate (4-5 pts)	t	n
Performance requirements (complexity of requirements such as response time and throughput)	4.08	0.90	71.9%	11.78	96
Operating environment compatibility (support for different HW/SW environments)	4.04	0.90	80.0%	11.25	95
Interface complexity (complexity based on the number of linked organizations)	4.03	0.79	76.8%	12.72	95
Scale (productivity adjustment based on project size)	3.95	0.72	78.1%	12.85	96
Security requirements (secure coding, vulnerability checks, encryption, etc.)	3.84	0.91	76.0%	9.09	96

4.6 SW 유형별 비교

응답자가 주로 참여하는 SW 유형별로 비교한 결과, 네 유형 모두 제한사항 공감도 3.73 이상, 필요성-반영수준 격차 1.15 이상으로 현행 산정체계 보완 필요성에 대한 방향은 일치하였다(Table 7). 기본형의 필요성(4.29)과 격차(1.49)가 가장 높게 나타나, 보완 필요성이 AI · 데이터 중심 사업에 국한되지 않고 일반 SW 개발을 포함한 전체 사업유형의 문제임을 보여준다. 모델형은 필요성 평균(3.68)이 낮아 격차도 작게 나타났으나, 이는 반영이 충분함을 의미하기보다 필요성 평가 자체가 낮았던 결과로 해석된다. 응답자별 평균 격차에 대한 일원분산분석 결과 유형 간 차이는 통계적으로 유의하지 않았으며($F(3, 92)=2.40$, $p=.073$), 유형별 표본 크기가 14～40명으로 상이하므로 유형 간 차이는 경향성을 확인하는 보조적 결과로 해석할 필요가 있다. 보고된 F값과 자유도로 계산한 효과크기(η^2)는 약 .073으로, 후속 연구에서는 유형별 표본을 균형 있게 확보하여 이 차이의 재현성을 확인할 필요가 있다.

〈Table 7〉 Comparison of Key Indicators by Software Type

SW Type (n)	Agreement with Limitations	Reflection Level	Necessity	Gap
Basic (40)	3.88	2.80	4.29	1.49
Data-centric (14)	3.98	2.69	4.00	1.31
Model-centric (22)	3.73	2.53	3.68	1.15
Varies by project (20)	3.80	2.57	4.05	1.48

www.kci.go.kr

4.7 개방형 응답 분석

추가로 고려할 비용 항목을 묻는 문항에는 10건, 개발비 산정 전반에 대한 의견 문항에는 14건의 응답이 있었으며, 내용을 범주화하면 다음과 같다. 첫째, AI・데이터 관련 비용으로 폐쇄망 환경에서의 AI 학습데이터 자체 구축 비용(외부 공개 데이터 활용 불가), 합성데이터 생성 및 라벨링 검수 인력 비용, GPU 서버 등 AI 학습 인프라 임차 또는 감가상각 비용, LLM 등 AI 기반 개발도구 사용료가 제시되었다. 특히 폐쇄망에서의 학습데이터 구축은 일반 SW 개발과 성격이 전혀 달라 가중단가 또는 별도 표준 산정기준이 필요하다는 의견과, AI 개발도구가 생산성을 높이지만 그 사용료를 반영할 항목이 없다는 의견이 있었다. 둘째, 사업 착수 전후의 비용으로 제안・검토 단계의 요구사항 분석 비용, 보안 USB・반출입 심의 등 자료 반출입 절차의 대기 공수, 시험평가 지원을 위한 현장 파견 인건비・출장비, 형상관리・품질보증 전담 인력 비용, 사업 종료 후 하자보수 기간의 무상 유지보수 인력 비용이 제시되었으며, SDRL 작성과 시험・검증에 투입되는 공수가 개발 공수와 비슷한 수준임에도 대가에는 거의 반영되지 않는다는 지적이 있었다. 셋째, 산정체계 전반에 대해서는 SW를 전문인력의 지식이 축적된 기술자산・창작물로 평가할 필요성, 정부 노임단가와 시장 임금의 격차로 고급 인력 투입이 어렵다는 점, 양산 및 후속 운용 단계에서 SW 개선이 사실상 무상으로 요구되는 관행, 클라우드・상용 라이선스 비용이 별도 항목으로 반영되지 않아 개발비를 줄여 충당하는 현실, 공수 증빙 자료 준비 부담과 증빙 기준의 간소화・표준화 필요성이 언급되었다. 개선 방식에 대해서는 기능점수 방식으로의 전환보다 현행 투입공수 방식을 유지하면서 국방 특수비용 항목을 추가하는 단계적 보완이 현실적이라는 의견이 있었다. 이러한 응답은 정량분석에서 격차가 크게 나타난 AI・데이터, 유지보수, 요구사항 변경, 시험・검증, 클라우드・라이선스 항목과 대체로 일관된 방향을 보이며, 정량 문항에 포함되지 않았던 착수 전 검토 비용, 반출입 대기 공수, 하자보수 기간 비용, AI 개발도구 사용료를 새로운 비용요소 후보로 제시한다. 다만 개방형 응답 수가 10건과 14건으로 제한적이므로, 해당 범주는 발생빈도의 추정치가 아니라 후속 검증이 필요한 비용요소 후보로 해석하였다.

V. 결론 및 논의

본 연구는 국방 SW 사업 참여기업을 대상으로 한 2단계 설문조사를 통해 현행 투입공수법 기반 개발비 산정체계에 대한 현장 인식을 진단하고 제도 보완의 우선순위를 도출하였다. 주요 결과를 요약하면 다음과 같다. 첫째, 현행 산정체계의 제한사항 8개 항목 모두에 대해 중립값보다 유의하게 높은 공감이 확인되었으며, 특히 자체 투자・기술자산의 보상 부족, 양산・후속 단계에서의 SW 가치 미인정, 국방 특수비용 미반영이 가장 큰 문제로 인식되었다. 둘째, 13개 비용 항목 모두에서 필

www.kci.go.kr

요성이 현재 반영수준을 유의하게 상회하였고 응답자의 97.9%에서 동일한 방향이 확인되었으며, 유지보수 · 성능개선과 AI 개발비용(테스트 · 검증, 데이터 구축, 모델 학습)의 격차가 가장 컸다. 이들 4개 항목은 IPA에서도 중점 개선 영역에 위치하였으며, 같은 영역에 위치한 '요구사항 변경 대응'과 '클라우드 · 라이선스'는 반영 수준이 기준축에 근접한 경계 항목으로 구분하였다. 셋째, AI 항목의 큰 격차는 필요성이 높아서가 아니라 반영수준이 유의하게 낮기 때문에 발생하였다. 넷째, 사업 특성에 따른 추가 보정과 5개 보정계수에 대해서는 전반적으로 긍정적인 수용성이 확인되었다.

분석 결과는 국방 SW 대가산정의 기본 방향에 대해 다음과 같은 점을 시사한다. 다만 이하의 내용은 앞서 요약한 조사 결과를 바탕으로 검토할 수 있는 방안을 저자가 제안하는 것이다. 특히 본조사는 투입공수법, 기능점수 방식 등 대안적 산정 방식에 대한 선호를 직접 측정하지 않았으므로, 특정 산정 방식에 대한 응답자 전체의 선호가 확인된 것은 아니다. 첫째, 산정 방식의 골격에 관해서는 투입공수법을 전면 대체하기보다 투입 인력 · 기간을 기본 산정 요소로 유지하면서 보완하는 접근을 우선 검토할 수 있다. 제한사항 가운데 '자체 투자 · 기술 자산 보상 부족'과 '고급 SW 인력 가치 반영 부족'에 대한 공감이 높았고, 개방형 응답에서 기능점수 방식으로의 전환보다 단계적 보완이 현실적이라는 의견이 일부 제시되었다. 이는 일부 응답자의 정성적 의견이므로 응답자 전체의 선호로 일반화할 수는 없으나, 현행 제도와의 연속성을 고려할 때, 기술 난이도 · 전문 인력 수준 · 보유 기술 자산을 기본 공수에 가산 · 조정하는 방식이 우선 검토될 수 있다. 둘째, 비용 요소의 구성에 관해서는 격차 분석과 IPA에서 일관되게 우선 항목으로 도출된 AI 개발비용(데이터 구축, 모델 학습, 테스트 · 검증)과 유지보수 · 성능개선 비용을 기존 항목의 요율 조정이 아닌 독립적인 비용 요소로 식별하는 방안을 검토할 수 있다. AI 항목의 격차가 필요성이 아닌 낮은 반영 수준에서 비롯되었다는 결과는 이들 비용이 현행 산정체계에서 충분히 반영되지 않는 것으로 인식되고 있음을 보여준다. 다만 이것이 비용을 계상할 항목 자체가 없기 때문인지, 항목은 있으나 실제 사업에서 충분히 인정되지 않기 때문인지는 본 조사로 구분되지 않으므로, 실제 산정내역서 분석을 통해 확인할 필요가 있다. 셋째, 시험 · 검증, 요구사항 변경, 클라우드 · 라이선스 비용은 반영수준의 편차가 커 항목의 신설보다 적용 기준의 일관성 확보가 과제이며, 이는 방산원가 기준과 SW사업 대가산정 가이드 간 차이에서 비롯되는 문제와 맞닿아 있다. 넷째, 보정계수에 관해서는 5개 계수 모두 수용성이 확인되었으나 보안성 요구수준은 세부 기준의 구체화가 필요하며, 개방형 응답에서 제시된 임무 중요도, 안전필수 요구, 시험검증 강도, 개발환경 제약, 요구사항 변동성 등은 추가 보정요소의 후보로 검토할 수 있다.

이상의 결과를 산정체계로 전환하면 <Table 8>과 같이 기본 개발비, 독립 비용 요소, 사업 특성 보정 및 수명주기 비용의 네 구성 요소로 정리할 수 있다. 다만 <표 8>은 인식 조사 결과를 바탕으로 산정체계가 갖추어야 할 구성 요소를 개념적으로 정리한 것으로, 실제 원가 자료로 추정 · 검증된 비용 산정 모형이 아니다. 따라서 이는 가산율이나 보정계수의 수준을 확정하는 계산식이 아니라 누락 비용을 식별하고 중복 산정을 방지하기 위한 개념적 구조이며, 구성 요소별 가산율과 보정계수의 수준 및 적용방식은 실제 사업비 자료를 통한 계량적 검증을 거쳐 결정되어야 한다.

www.kci.go.kr

〈Table 8〉 Principles for Designing the Defense Software Development Cost Estimation Framework

Component	Main Estimation Targets	Evidence from This Study	Application Principle
Base development cost	Personnel and duration, overhead, technical fee, direct expenses	Agreement with 8 limitations (3.71-4.00); open-ended responses favoring phased improvement (qualitative evidence)	Retain man-hours as the baseline for cost estimation while separately assessing the contributions of highly experienced personnel and technological assets
Independent cost elements	AI data preparation, model training and validation, software testing and quality assurance (QA), cloud services and software licensing, and requirements changes	Necessity significantly exceeded reflection level for all 13 items	Record these costs as separate line items in the cost estimate rather than incorporating them into the baseline man-hour calculations
Project characteristic adjustment	Project size, interface complexity, performance requirements, system compatibility, security requirements, mission criticality, and other relevant factor	Mean adequacy of 5 adjustment factors: 3.84-4.08	Standardize the conditions for application, classification levels, and supporting documentation to prevent double counting
Life-cycle cost (post-fielding)	Post-fielding maintenance, retraining, performance improvement, response to vulnerabilities and data changes	Maintenance & performance improvement gap: +1.78	Estimate operating costs separately from development costs, and establish the operational period and update cycle during the planning stage to ensure separate budget allocation

제도 적용은 기획, 계약, 수행・운영의 세 단계로 구분할 필요가 있다. 기획단계에서는 비용요소별 작업분류체계와 산정근거를 제시하고, 계약단계에서는 요구사항 변경, AI 성능목표 미달 및 데이터 추가확보를 예산 조정의 발동조건으로 정해야 한다. 수행・운영단계에서는 실제 투입비용과 변경이력을 축적하여 보정계수와 표준단가를 주기적으로 보정해야 한다. 이러한 절차가 병행되어야 보정계수가 일률적 인상률로 운용되거나 동일 비용이 기본공수와 별도 항목에 중복 반영되는 문제를 줄일 수 있다.

<Table 8>의 구성 요소 중 기본 개발비, 독립 비용 요소, 사업 특성 보정은 체계개발 단계의 개발비에 해당하며, 수명주기 비용은 전력화 이후 발생하는 비용으로서 개념・예산・계약의 측면에서 개발비와 구분된다. 개념상 개발비는 요구사항을 충족하는 SW를 개발하고 시험평가를 거쳐 납품하기까지의 비용이며, 수명주기 비용은 전력화 이후 운용 과정에서 발생하는 결함 수정, 운용 환경 변화 대응, 보안 취약점 조치, AI 모델 재학습과 성능개선 비용을 의미한다. 예산 측면에서 개발비는 방위력개선비의 연구개발 예산으로 편성되는 반면, 전력화 이후의 유지보수는 주로 전력운영비로, 기능 추가나 성능개량

www.kci.go.kr

은 별도의 성능개량사업으로 편성되어 재원과 편성 주체가 다르다. 계약 측면에서도 개발비는 체계개발 계약에서 산정·정산되고, 수명주기 비용은 별도의 유지보수 또는 후속지원 계약의 대상이 되며, 그 사이의 하자보수 기간에는 개발계약상 하자보수 의무에 따라 별도 대가 없이 결함 수정이 이루어진다. 개방형 응답에서 제기된 하자보수 기간의 무상 유지보수 요구나 양산·후속 단계의 무상 개선 요구는 이 경계가 계약에서 명확히 구분되지 않을 때 발생하는 문제로 볼 수 있다. 따라서 본 연구에서 제시한 수명주기 비용은 개발비에 가산하는 항목이 아니라, 개발비와 구분하되 체계개발 기획 시점에 그 소요를 함께 식별하여 후속 예산과 계약으로 연결하는 비용으로 정의하며, 특히 하자보수의 범위(결함 수정)와 유지보수·성능개선의 범위(환경 변화 대응, 재학습, 기능 개선)를 계약에서 구분할 필요가 있다.

아울러 제안 체계가 현행 제도에 접목되기 위해서는 무기체계 사업에 적용되는 방산원가 기준과 공공 SW 사업에 적용되는 「SW사업 대가산정 가이드」의 차이를 고려해야 한다. 두 기준의 주요 차이와 본 연구의 제안 체계와의 연결 관계를 정리하면 <Table 9>와 같다.

〈Table 9〉 Comparison of Current Cost Estimation Standards and Directions for Integration into the Proposed Framework

Category	Defense Industry Cost Standard	SW Project Cost Estimation Guide	Link to Proposed Framework
Scope of application	Defense cost-subject items (weapon system R&D, production, etc.)	Public-sector SW projects (informatization projects, etc.)	Since the applicable standard for the same SW development activity varies depending on the legal classification of the project, the proposed framework is designed to identify cost elements that can be applied under both standards
Cost structure	Manufacturing cost structure: materials, labor, expenses, general administrative costs, profit, etc.	Direct labor, overhead, technical fee, direct expenses (man-hour method) or function point-based development cost	Base development cost uses each standard's labor cost structure; independent cost elements are mapped to the corresponding cost categories of each standard
Technology compensation & profit	Profit calculated according to the compensation structure set by the cost accounting rules	Technology compensation reflected via the technical fee (20-40% of direct labor plus overhead)	Examine how the contributions of highly experienced personnel and technological assets can be reflected separately under each standard through profit compensation provisions and differentiated technical fee rates, respectively
Project characteristic adjustment	No explicit adjustment system for SW characteristics	The function point method incorporates adjustment factors for project size, interface complexity, performance requirements,	Extend the adjustment factors specified in the Guide to defense projects while incorporating defense-specific considerations, such as mission criticality and verification

www.kci.go.kr

Category	Defense Industry Cost Standard	SW Project Cost Estimation Guide	Link to Proposed Framework
		operating environment compatibility, and security requirements	intensity
Operational-phase SW cost	Focused on R&D and production costs; limited criteria for estimating post-fielding SW improvement costs	Provides separate criteria for estimating software maintenance and operating costs	Estimate life-cycle costs separately from development costs and examine the applicability of the Guide's maintenance cost estimation methodology to defense projects

즉 제안 체계는 기존의 두 기준을 대체하는 별도의 산정기준이 아니라 두 기준이 공통으로 활용할 수 있는 비용 요소 식별 구조로서, 방산원가 기준이 적용되는 사업에서는 원가계산서의 비목에, 대가산정 가이드가 적용되는 사업에서는 산정내역서의 항목에 각각 대응시키는 방식으로 접목될 수 있다. 앞서 시험・검증, 요구사항 변경, 클라우드・라이선스 비용에서 확인된 반영 수준의 편차를 줄이기 위해서도 동일한 비용 요소가 두 기준에서 어떤 비목으로 인정되는지를 대응표로 명시하는 작업이 우선될 필요가 있다.

본 연구의 학술적 기여는 국방 SW 개발비 산정이라는 정책 현안에 대해 정성・정량 조사를 순차적으로 결합하여 현장 인식을 체계적으로 계량화하고, 격차 분석과 IPA, 응답자 단위 분석과 클러스터 비교를 병행하여 제도 보완의 우선순위를 교차 검증하였다는 데 있다. 그럼에도 다음과 같은 한계가 있다. 첫째, 본 조사는 인식조사이므로 구체적인 가산율이나 보정계수 수준을 확정하는 근거로 단독 활용하기에는 제한이 있으며, 본 연구는 산정체계의 기본 방향을 진단하는 데 초점을 두어 제도・규정의 개정 방안은 다루지 않았다. 둘째, 예비조사는 중소기업, 본조사는 기업 실무자 중심으로 이루어져 대기업・체계업체 및 발주기관의 관점이 충분히 반영되지 못하였다. 또한 본조사에서 응답자의 소속 기업 규모, 직무, 경력 등 배경 변수를 수집하지 않아 응답자 특성에 따른 인식 차이를 분석하지 못하였다. 특히 응답자 전원이 공급자 측 실무자이므로 반영 수준은 낮게, 필요성은 높게 평가하는 방향의 인식 편향이 있을 수 있으며, 이에 따라 격차의 절대적 크기가 과대 추정되었을 가능성이 있다. 다만 AI・비AI 클러스터 비교나 항목 간 우선순위와 같은 응답자 내 상대 비교는 이러한 편향이 항목 전반에 공통으로 작용한다면 상대적으로 영향을 덜 받는다. 이러한 편향을 검증하기 위해서는 발주기관(방위사업청, 각 군)과 체계업체를 대상으로 동일 문항의 조사를 실시하여 공급자와 발주자 간 인식 차이를 비교하고, 실제 계약 산정내역서와 집행 자료에서 비용 항목별 인정 여부와 금액을 확인하는 후속 검증이 필요하다. 셋째, 문항별 무응답으로 유효응답 수에 차이가 있고, SW 유형별 표본 크기가 상이하여 유형 간 비교의 일반화에는 신중할 필요가 있다. 넷째, 측정문항이 개별 비용요소를 평가하는 형성적 지표로 설계되어 내적일관성에 기반한 신뢰도 검증을 적용하기 어려우며, 향후 척도 개발 시 반영적 지표를 병행할 필요가 있다. 다섯째, 필요성과 반영수준을 동일 응답자가 같은 시점에 평가한 자기보고 자료이므로 공통방법편의와 응답 정박효과의 가능

www.kci.go.kr

성이 있다. 여섯째, IPA의 사분면 배치는 표본 내부 기준에 의존하므로 기준축 변화에 따른 분류 안정성을 추가로 검증해야 한다. 향후 연구에서는 대기업·체계업체와 발주기관을 대상으로 조사 범위를 확장하여 관점을 보완하고, 실제 사업비 자료를 활용하여 본 연구에서 도출된 우선 비용요소의 가산율과 보정계수의 수준을 계량적으로 검증할 필요가 있다. 이때 발주기관·체계업체·전문기업의 실제 계약 및 집행자료를 연계하고, 사업 종료 후 예측비용과 실제비용의 오차를 비교하여 비용항목과 보정계수의 예측타당성을 검증할 필요가 있다.

Acknowledgements

이 논문은 2026년도 국방기술진흥연구소 재원을 지원받아 수행한 연구 결과임(No. R26TA0183 7165, 국방 소프트웨어 사업 개발비 추정 방안 연구)

This research was supported by the Korea Research Institute for Defense Technology Planning and Advancement (KRIT) in 2026 under the project titled "A Study on Development Cost Estimation Methods for Defense Software Projects" (No. R26TA01837165).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Author Contributions

Conceptualization: LJS and CEB, Literature Review: LJS and HHH, Resources: CEB, Data Curation: HHH and CEB, Formal Analysis: HHH, Investigation and Methodology: LJS, Writing (Original Draft): LJS and HHH, Project Administration and Supervision: LJS.

www.kci.go.kr

Reference

Albrecht, A. J. (1979). Measuring application development productivity. In Proceedings of the Joint SHARE/GUIDE/IBM Application Development Symposium (pp. 83–92).

Boehm, B. W., Abts, C., Brown, A. W., Chulani, S., Clark, B. K., Horowitz, E., Madachy, R., Reifer, D. J., & Steece, B. (2000). Software cost estimation with COCOMO II. Prentice Hall.

Bollen, K. A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. Psychological Bulletin, 110(2), 305–314. https://doi.org/10.1037/0033-2909.110.2.305

Clark, B., & Madachy, R. (Eds.). (2015). Software cost estimation metrics manual for defense systems. Software Metrics Inc. https://softwarecost.org/

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

Defense Acquisition Program Administration. (2024). Manual for weapon system software development and management (DAPA Manual No. 2024-6).

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. Journal of Marketing Research, 38(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Galorath, D. D., & Evans, M. W. (2006). Software sizing, estimation, and risk management: When performance is measured performance improves. Auerbach Publications.

Jørgensen, M., & Shepperd, M. (2007). A systematic review of software development cost estimation studies. IEEE Transactions on Software Engineering, 33(1), 33–53. https://doi.org/10.1109/TSE.2007.256943

Korea Artificial Intelligence and Software Industry Association. (2025). SGuide to software project cost estimation (2025 revised edition). https://www.swai.or.kr/

Kwon, T. Y. (2010). Development of a cost estimation model for embedded software in weapon systems considering the defense environment, Master's thesis, Korea Advanced Institute of Science and Technology.

Lee, B. E., & Kang, S. J. (2005). A development cost estimation at initial phase for military software using backfiring approach. Journal of the Korea Information Processing Society D, 12D(5), 737–744.

Martilla, J. A., & James, J. C. (1977). Importance-performance analysis. Journal of Marketing, 41(1), 77–79. https://doi.org/10.2307/1250495

Ministry of National Defense. (2026). Regulations on cost accounting for defense cost items (Ministry

www.kci.go.kr

of National Defense Ordinance No. 1202). National Law Information Center.

Park, R. E. (1988). The central equations of the PRICE software cost model [Conference presentation]. 4th COCOMO Users' Group Meeting, Software Engineering Institute, Carnegie Mellon University.

Rosa, W., Clark, B. K., Madachy, R., & Boehm, B. W. (2022). Empirical effort and schedule estimation models for agile processes in the US DoD. IEEE Transactions on Software Engineering, 48(8), 3117–3130. https://doi.org/10.1109/TSE.2021.3080666

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. Advances in Neural Information Processing Systems, 28, 2503–2511.

Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

U.S. Department of Defense. (2020, October 2). Operation of the software acquisition pathway (DoD Instruction 5000.87).

원 고 접 수 일 2026년 09월 21일
원 고 수 정 일 2026년 10월 01일
게 재 확 정 일 2026년 10월 02일

www.kci.go.kr

IDAP 광운대학교 방위사업연구소 INSTITUTE OF DEFENSE ACQUISITION PROGRAM

JAMS JOURNAL OF ADVANCES IN MILITARY STUDIES

2026, Vol. 9, No. 2, 119-141.

https://doi.org/10.37944/advmil.9.2.202610.006

국방 소프트웨어 개발비 산정체계 개선방안 연구

이지선* · 황현호 · 최은비*****

국문초록

국방분야에서 소프트웨어(SW)와 인공지능(AI) 기반 첨단체계로 전환되고 SW가 무기체계 성능의 핵심 요소로 부상하고 있으나, 국방 SW 개발비는 여전히 투입공수와 인건비 중심으로 산정되어 사업 특성과 기술적 난이도의 반영이 미흡하다. 또한 방산원가 기준과「SW사업 대가산정 가이드」간 차이에 따른 제도적 문제도 제기되고 있다. 본 연구는 현행 산정체계에 대한 실무자 인식을 분석하고 개선방향을 제시하였다. 국방 SW 관련 중소기업 35개사의 예비조사로 문항을 도출한 후 실무자 96명을 대상으로 5점 척도 설문을 실시하였다. 분석 결과, 현행 제도의 제한사항 8개 항목은 모두 중립값보다 유의하게 높았으며(평균 3.71~4.00, $p<.001$), 13개 비용 항목에서도 필요성이 현행 반영수준을 유의하게 상회하였다($d=0.36\sim1.33$). 특히 유지보수 · 성능개선, AI 테스트 · 검증, 데이터 구축 및 모델 학습의 격차가 가장 컸다(+1.64~+1.78). AI 항목은 비AI 항목과 필요성에는 차이가 없었으나 반영수준은 유의하게 낮았다. 사업 특성별 추가 보정 필요성(4.12)과 5개 보정계수의 적정성(3.84~4.08)도 긍정적으로 평가되었다. 따라서 투입공수를 기본요소로 유지하되 AI 개발 및 수명주기 비용을 독립적으로 식별하고 사업 특성별 보정계수를 적용하는 방향으로 국방 SW 개발비 산정체계를 개선할 필요가 있다. 이러한 결과를 바탕으로 투입공수를 기본요소로 유지하되 AI 개발 및 수명주기 비용을 독립적으로 식별하고 사업 특성별 보정계수를 적용하는 방향을 국방 SW 개발비 산정체계의 개선방안으로 검토할 수 있으며, 실제 원가 자료를 통한 검증이 필요한 개념적 틀을 제시하였다.

주제어 : 국방 소프트웨어 개발비, 소프트웨어 대가산정, 인공지능 개발비, 수명주기 비용, 사업 특성별 보정계수

* (제1저자, 교신저자) (사)한국국방기술학회, jisun0319@gmail.com, https://orcid.org/0000-0003-2345-7698

** (공동저자) 한양사이버대학교 국방융합기술학과,1250002@hycu.ac.kr, https://orcid.org/0009-0001-4065-0943

*** (공동저자) 국방기술진흥연구소 비용분석팀, choieunbi@krit.re.kr

www.kci.go.kr